Advanced Microeconomic Theory: Economics 521b
Spring 2011
Juuso Välimäki

Mechanism Design: Basic Concepts

The setup is similar to that of a Bayesian game. The ingredients are:

1. Set of players, $i \in \{1, 2, ..., N\}$.

2. Sets of possible types for all players, Θ_i for $i \in \{1, 2, ..., N\}$. Let $\theta_i \in \Theta_i$ denote a typical type of player i.

3. Prior probability distribution $p(\theta)$ on $\Theta = \Theta_1 \times \cdots \times \Theta_N$.

4. Set of possible collective decisions Q.

5. Sets of possible actions (messages), $M_1, ..., M_N$ to the players and decision and transfer rules, $q : M \to \Delta(Q)$ and $t_i : M \to \mathbb{R}$. We let $\phi(m) = (q(m), t(m))$. All of these are to be chosen by the mechanism designer.

6. Utility functions $u_i : \Theta \times Q \times \mathbb{R} \to \mathbb{R}$,for $i \in \{1, 2, ..., N\}$ and a payoff to the mechanism designer, $v : \Theta \times Q \times \mathbb{R} \to \mathbb{R}$. In much of these notes, we concentrate on quasilinear payoffs:

$$u_i(\theta, q, t) = u_i(\theta, q) - t, \; v(\theta, q, t) = v(\theta, q) + t.$$

7. Once the sets M_i and the allocation rule ϕ have been fixed, the players play the Bayesian game with action sets M_i and payoff functions $\widetilde{u}_i : M \times \Theta \to \mathbf{R}$, where $u_i(m, \theta) = u_i(\theta, q(m), t(m))$.

We call the pair $\Gamma = (M, \phi)$ a mechanism. We will consider dominant strategy equilibria in these Bayesian games as well as Bayesian Nash equilibria. Whichever solution concept we adopt, we call $\phi(m) = (q(m), t(m))$ the outcome according

to that solution concept. We denote the strategy of player i in the game induced by Γ with

$$\sigma_i : \Theta_i \to \Delta(M_i).$$

The first three items above are similar to those in Bayesian games. The differences are introduced in items 4-6. Observe from 6 that the mechanism designer has a payoff in these games. Finally, the big difference to regular Bayesian games is that the action sets and the outcome functions are chosen by the mechanism designer. Since there are infinitely many possible choices for the sets M_i as well as for the functions ϕ, the task of analyzing games of this type may seem at first impossible. The following observation, however, makes the task much simpler and allows us to get started with the analysis. A similar theorem with completely analogous proof applies to dominant strategy solutions

Proposition 1. *(Revelation Principle for BNE) For any mechanism* $\Gamma = (M, \phi)$, *and any* BNE σ *of* Γ, *there is another mechanism* $\Gamma' = (M', \phi')$ *and an equilibrium* σ' *of* Γ' *such that for every profile of types* θ, *the (possibly random) equilibrium outcome in* Γ *under* σ *coincides with the outcome under* σ' *in* Γ'. *Furthermore, we can take* $M'_i = \Theta_i$ *and* $\sigma'_i(\theta_i) = \theta_i$ *for all* θ_i.

Proof. We set

$$\phi'(\theta) = \phi(\sigma(\theta)).$$

If

$$\sigma'(\theta) = \theta$$

is not a BNE of the game induced by Γ', then there is an i and $\theta_i, \theta'_i \in \Theta_i$ with

$$\Sigma_{\theta_{-i}} u_i(\theta, \phi'(\theta'_i, \theta_{-i})) p(\theta_{-i} | \theta_i) > \Sigma_{\theta_{-i}} u_i(\theta, \phi'(\theta_i, \theta_{-i})) p(\theta_{-i} | \theta_i)$$

or

$$\Sigma_{\theta_{-i}} u_i(\theta, \phi(\sigma_i(\theta'_i), \sigma_{-i}(\theta_{-i}))) p(\theta_{-i} | \theta_i) > \Sigma_{\theta_{-i}} u_i(\theta, \phi(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i}))) p(\theta_{-i} | \theta_i).$$

But this contradicts the optimality of σ_i in the game induced by Γ. $\square$

Observe the reduction in complexity afforded by this result. We call mechanisms with $M_i = \Theta_i$ for all i *Direct Mechanisms*. Strategies satisfying $\sigma_i(\theta_i) = \theta_i$ are called truthful. Revelation Principle gives a sense in which there is no loss of generality in considering only truthful equilibria of direct mechanisms.

Comments:

1. Notice the commitment possibility by the designer. She cannot change ϕ after observing messages θ_i.

2. Γ' may have other equilibria than σ'. How are we to know that σ' is the relevant one? The others may be completely different and even superior to the players. Part of mechanism design literature deals with this question, and requires that σ' be the only equilibrium. This literature goes under the names implementation and full implementation.

3. Notice that we assume simultaneity in the announcement stage and all announcements in a single stage. What if all information is not available at the outset?

4. All information is treated the same, i.e. payoff relevant and higher order information.

5. Designer has no private information. Otherwise we have informed principal problems.

6. With multiple principals, revelation principle may fail.

Characterizing Incentive Compatible Allocations

We review the basic charcterization of incentive compatible mechanisms in a special class of environments that turns out to be the most useful for applications. We assume that the types of the players are independently distributed and that the payoff functions take a simple form. Furthermore, we assume that we are in the private values case where the utility of player i depends on θ_j only through the allocation.

Linear Environments

We turn next to the question of what kinds of outcomes (q, t) are compatible with Bayesian equilibrium. We will also specialize the scope of the models a little by requiring that the utilities take the following linear form:

$$u_i(\theta, q, t) = \theta_i v_i(q) - t_i(\theta).$$

We also assume that for all i, $\theta_i \in \left[\underline{\theta}_i, \overline{\theta}_i\right]$ and that the distribution on the types is twice continuously differentiable. Furthermore, we assume that for $i \neq j$, $\theta_i \perp \theta_j$. We look for a characterization of the truthful equilibria of direct mechanisms. Denote an arbitrary announcement by player i by $\widehat{\theta}_i$, and denote by $T_i\left(\widehat{\theta}_i, \theta_i\right)$ the expected transfer given true type θ_i and announcement $\widehat{\theta}_i$. Notice that by our assumption of independence of types, we can write

$$T_i\left(\widehat{\theta}_i, \theta_i\right) = T_i\left(\widehat{\theta}_i\right).$$

Similarly, let $V_i\left(\widehat{\theta}_i\right) = E_{\theta_{-i}}\left[v_i\left(q\left(\widehat{\theta}_i, \theta_{-i}\right)\right)\right]$. Because of the linearity of the utilities, we may now write the expected payoff from announcing $\widehat{\theta}_i$ while of type θ_i.

$$E_{\theta_{-i}}\left[u_i\left(q\left(\widehat{\theta}_i, \theta_{-i}\right), \theta\right)\right] = \theta_i V_i\left(\widehat{\theta}_i\right) + T_i\left(\widehat{\theta}_i\right),$$

Define also the following functions for $i = 1, ..., N$:

$$U_i\left(\theta_i, \widehat{\theta}_i\right) = \theta_i V_i\left(\widehat{\theta}_i\right) + T_i\left(\widehat{\theta}_i\right).$$

Now we are in a position to characterize the outcomes that are achievable in Bayesian Nash equilibria of mechanism design problems.

Proposition 2. *((Bayesian incentive compatibility of outcomes)*
$(q(\theta), t(\theta))$ *is Bayesian incentive compatible if and only if for all* $i = 1, ..., N$,
i) $V_i\left(\widehat{\theta}_i\right)$ *is nondecreasing.*
ii) $U_i(\theta_i) = U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} V_i(s)\, ds$ *for all* θ_i.

Proof. This is in e.g. MWG p888-889. □

This is a really remarkable result. It tells us immediately that $U_i(\underline{\theta}_i)$ and $q(\theta)$ fixes the expected payoffs of all players in any truthful equilibrium of any direct mechanism. In a sense it uses the envelope theorem to eliminate the transfers from the problem. Therefore also the expected transfers are fixed by $q(\theta)$ and $U_i(\underline{\theta}_i$ In first year micro lectures, you saw how this result can be used to show revenue equivalence of various auction forms. In these notes, we shall see how to use this information for deducing properties of social surplus maximizing allocation schemes.

Aside: Optimal Mechanism with a Single Agent

Payoff to agent from allocation $(q(\theta'), t(\theta'))$ when true type is θ is given by:

$$U(\theta, \theta') = u(q(\theta'), \theta) - t(\theta').$$

Agent announces θ' to maximize welfare.

First order condition for truthful announcements, i.e. $\theta' = \theta$:

$$\frac{\partial u(q(\theta'), \theta)}{\partial q} q'(\theta') = t'(\theta') \text{ at } \theta' = \theta.$$

Thus

$$\frac{\partial u(q(\theta), \theta)}{\partial q} q'(\theta) = t'(\theta). \tag{1}$$

Let $U(\theta) = U(\theta, \theta)$.

Then (**??**) implies that:

$$U'(\theta) = \frac{\partial u(q(\theta), \theta)}{\partial q} q'(\theta) - t'(\theta) + \frac{\partial u(q(\theta), \theta)}{\partial \theta} = \frac{\partial u(q(\theta), \theta)}{\partial \theta}.$$

Hence:

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q(s), s)\, ds.$$

Second order condition for truthtelling at $\theta' = \theta$:

$$\frac{\partial^2 u(q(\theta), \theta)}{\partial q^2} (q'(\theta))^2 + \frac{\partial u(q(\theta), \theta)}{\partial q} q''(\theta) - t''(\theta) \leq 0.$$

Differentiate **??** with respect to θ to get:

$$\frac{\partial^2 u(q(\theta), \theta)}{\partial q^2} (q'(\theta))^2 + \frac{\partial u(q(\theta), \theta)}{\partial q} q''(\theta) + \frac{\partial^2 u(q(\theta'), \theta)}{\partial q \partial \theta} q'(\theta) = t''(\theta).$$

Combining these last two yields:

$$-\frac{\partial^2 u(q(\theta'), \theta)}{\partial q \partial \theta} q'(\theta) \leq 0.$$

By Spence-Mirrlees single crossing, this implies:

$$q'(\theta^{)} \geq 0.$$

Hence the above shows that

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q(s), s)\, ds \tag{2}$$

and

$$q'(\theta) \geq 0 \tag{3}$$

are necessary for truthtelling (i.e. necessary conditions for implementability).

To see sufficiency, observe that local first and second order conditions are satisfied by any allocation that satisfies the above. The only remaining question is

whether there can be another local maximum $\theta' = \theta^* \neq \theta$ for a mechanism where **??** and **??** hold.

To see that this is not possible, note from **??** that

$$\frac{\partial U(\theta, \theta')}{\partial \theta'} = [\frac{\partial u(q(\theta'), \theta)}{\partial q} - \frac{\partial u(q(\theta'), \theta')}{\partial q}] q'(\theta').$$

Hence by **??**, the right hand side of the above equation has the same sign as $(\theta - \theta')$ by the Spence-Mirrlees single crossing property. Thus $\theta' = \theta$ is the global maximizer

At this point, we know that there is no loss of generality in restricting attention to mechanisms satisfying **??** and **??**.

Payoff to principal:

$$V(\theta) = t(\theta) + v(q(\theta), \theta).$$

Let

$$S(q, \theta) = v(q, \theta) + u(q, \theta).$$

Then

$$V(\theta) = S(q, \theta) - U(\theta)$$

and

$$EV(\theta) = \int_{\underline{\theta}}^{\overline{\theta}} [S(q, \theta) - U(\theta)] f(\theta) d\theta =$$

$$\int_{\underline{\theta}}^{\overline{\theta}} [S(q, \theta) - U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q(s), s) ds] f(\theta) d\theta =$$

$$\int_{\underline{\theta}}^{\overline{\theta}} S(q, \theta) f(\theta) d\theta - \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\theta} u_\theta(q(s), s) ds f(\theta) d\theta =$$

$$\int_{\underline{\theta}}^{\overline{\theta}} S(q, \theta) f(\theta) d\theta - \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\theta} u_\theta(q(s), s) f(\theta) ds d\theta =$$

$$\int_{\underline{\theta}}^{\overline{\theta}} S(q, \theta) f(\theta) d\theta - \int_{\underline{\theta}}^{\overline{\theta}} [\int_{s}^{\overline{\theta}} f(\theta) d\theta] u_\theta(q(s), s) ds =$$

$$\int_{\underline{\theta}}^{\overline{\theta}} S(q,\theta) f(\theta)\, d\theta - \int_{\underline{\theta}}^{\overline{\theta}} (1-F(s))\, u_\theta(q(s), s)\, ds =$$

$$\int_{\underline{\theta}}^{\overline{\theta}} [S(q,\theta) - \frac{(1-F(\theta))}{f(\theta)} u_\theta(q(\theta), \theta)] f(\theta)\, d\theta.$$

Hence the problem of the principal is to

$$\max_{q(\theta)} \int_{\underline{\theta}}^{\overline{\theta}} [S(q,\theta) - \frac{(1-F(\theta))}{f(\theta)} u_\theta(q(\theta), \theta)] f(\theta)\, d\theta$$

$$\text{s.t. } q'(\theta) \geq 0.$$

How to solve: Try first by solving for optimal q for each θ separately. If for this $q(\theta)$, monotonicity is satisfied, then the problem is solved. If not, then optimal ocntrol theory is needed.

Efficient Mechanisms

For the case of private values and quasilinear payoff functions, there exists a simple mechanism that implements the socially efficient decision rule in dominant strategies. Let

$$q^*(\theta) \in \arg\max_{q \in Q} \sum_{i=1}^{N} u_i(\theta, q) = \arg\max_{q \in Q} \sum_{i=1}^{N} u_i(\theta_i, q),$$

where the last equality follows from the private values assumption.

Consider the direct mechanism with choice rule $q^*(\theta)$ and with transfer function

$$t_i\left(\widehat{\theta}\right) = -\sum_{j \neq i} u_j(\widehat{\theta}_j, q^*\left(\widehat{\theta}\right)) + \tau_i\left(\widehat{\theta}_{-i}\right).$$

It is easy to see that for an arbitrary announcement $\widehat{\theta}_{-i}$ by the other players, $\widehat{\theta}_i = \theta_i$ is a dominant strategy for an arbitrary function τ_i. The mechanism $(q^*(\theta), t(\theta))$ as above is called a *Vickrey-Clarke-Groves* mechanism (VCG mechanism). Green and Laffont (1977, 1979) show that if there are no restrictions on the domain of

preferences for the players, then VCG mechanisms are the only nmechanisms that make truthful revelation a dominant strategy for the efficient transfer rule $q^*(\theta)$.

A particularly useful VCG mechanism is the pivotal mechanism or the externality mechanism, where

$$\tau_i\left(\widehat{\theta}_{-i}\right) = \max_{q\in Q}\sum_{i\neq j} u_i(\widehat{\theta}_i, q).$$

If we denote the maximized surplus for a coalition $\mathcal{S}$ of players by

$$W_{\mathcal{S}}(\theta) = \max_{q\in Q}\sum_{i\in\mathcal{S}} u_i(\widehat{\theta}_i, q),$$

then we see that the utility in the pivotal mechanism going to player i is given by

$$\max_{q\in Q}\sum_{i=1}^{N} u_i(\theta, q) - \max_{q\in Q}\sum_{j\neq i} u_j(\widehat{\theta}_j, q) = W_{\{1,\ldots,N\}}(\theta) - W_{\{1,\ldots,N\}\setminus\{i\}}(\theta) := M_i(\theta),$$

i.e. in the pivotal mechanism, each player gets her marginal contribution to the social welfare as her payoff.

Budget Balance

An unfortunate feature of the VCG mechanism is that it is not budget balanced (BB). In other words, it is not necessarily the case that

$$\sum_{i=1}^{N} t_i(\theta) = 0$$

There are two senses in which the budget can be balanced: *ex ante, and ex post.* The differences here reflect what is known about the type vector. Ex post sense requires budget balance for all θ. Ex ante requires budget balance on average. Clearly, ex post budget balance is more restrictive than ex ante budget balance. In the quasilinear case with independent types, it is an easy exercise to show that the two notions of budget balance are equivalent in the sense that for each ex ante BB mechanism, there exists another mechanism that is ex post BB, where the social

decision rule is as in the original decision rule and the expected transfers across the two mechanisms are equal. Hence we will use ex post BB in what follows as our definition of BB.

AGV mechanism given by $q^*(\theta)$, and transfers

$$t_i\left(\widehat{\theta}\right) = -E_{\theta_{-i}} \sum_{j \neq i} u_j(\theta_j, q^*\left(\widehat{\theta}_i, \theta_{-i}\right)) + \tau_i\left(\widehat{\theta}_{-i}\right).$$

It is again easy to see that $\widehat{\theta}_i = \theta_i$ is an optimal announcement (not a dominant strategy though, why?).

Let

$$\mathcal{E}_i\left(\widehat{\theta}_i\right) = E_{\theta_{-i}} \sum_{j \neq i} u_j(\theta_j, q^*\left(\widehat{\theta}_i, \theta_{-i}\right)),$$

and set

$$\tau_i\left(\widehat{\theta}_{-i}\right) = \frac{\sum_{j \neq i} \mathcal{E}_j\left(\widehat{\theta}_j\right)}{N-1}.$$

Then the mechanism is ex post BB.

Notice that in the AGV mechanism, the players commit to participate in the mechanism at the ex ante stage. In this sense, we could view the situation as an example of moral hazard. The mechanism is agreed upon at a stage where the players and the designer have symmetric information. In the next section, we investigate what happens when the participation decisions are at the interim stage (i.e. when the players know their types).

IC, BB and Participation

The previous section showed that when one uses BNE as the solution concept budget balance can be achieved if one relaxed the participation constraint to be satisfied only in the ex ante sense. One should note, though, that as long as one insists on dominant strategy implementation, VCG are the only mechanisms that implement the efficient choice rule. In this section, we show in a public goods provision example that no VCG mechanism achieves budget balance. Hence we conclude that it is impossible to achieve BB and efficiency in a dominat strategy

mechanism. We also show that if one insists on IR or participation constraint being satisfied at the interim stage, then there are no efficient Bayesian mechanisms with BB either.

We discuss next the issue of individual rationality or participation in a mechanism. To fix ideas, we consider a game of public goods provision. A project can be completed at cost $c > 0$ or alternatively it may remain uncompleted. at zero cost. Therefore $q \in \{0, 1\}$. Assume that $u_i(\theta, 0) = 0$ for all i and $u_i(1, \theta) = \theta_i$ for all i, and θ_i. Then the efficient decision rule is to have

$$q^*\left(\widehat{\theta}\right) = \begin{cases} 1 \text{ if } \Sigma_i \widehat{\theta}_i \geq c. \\ 0 \text{ otherwise.} \end{cases}$$

Consider the pivot mechanism for this problem:

$$t_i\left(\widehat{\theta}\right) = \begin{cases} \underline{\theta} q^*\left(\underline{\theta}, \widehat{\theta}_{-i}\right) + c - \Sigma_{j \neq i} \widehat{\theta}_j \text{ if } \Sigma_i \widehat{\theta}_i \geq c \geq \Sigma_{j \neq i} \widehat{\theta}_j, \\ 0 \text{ otherwise.} \end{cases}$$

As a very easy exercise, verify that all players have a dominant strategy to report $\widehat{\theta}_i = \theta_i$. Suppose next that all players have the opportunity to opt out from the mechanism and get a utility of 0. Then it is immediate from the revenue equivalence and the transfer function that the largest transfers consistent with participation by all types are given by the pivot mechanism transfers. If we understand budget balance to mean that the sum of transfers covers the cost c, then we can simply evaluate the expected payments in the pivot mechanism and ask whether the cost c is covered. It is straightforward to show that cost is covered only in cases where $N\underline{\theta} > c$ or $c < N\overline{\theta}$. In all other cases, there is a deficit. The computation of this is left as an exercise.

The same line of argument can be used in e.g. the bilateral trade model. Construct the pivot mechanism. Argue that it has the largest transfers, and show that the pivot mechanism transfers result in a deficit.